УДК 616.07:681.518.54(045)

И. Г. Прокопенко, д-р техн. наук, проф.,
А. И. Чурина, канд. техн. наук, доц.

ИССЛЕДОВАНИЕ ЭФФЕКТИВНОСТИ СТАТИСТИЧЕСКИХ КРИТЕРИЕВ В ЗАДАЧАХ ОБРАБОТКИ БИОМЕДИЦИНСКИХ ДАННЫХ

Институт электроники и систем управления НАУ, e-mail: iesy@nau.edu.ua

Рассмотрена задача сравнительного анализа эффективности трех алгоритмов принятия решений о равенстве распределений вероятностей двух выборок – критерия Стьюдента, критерия Фишера и двухстороннего рангового критерия. Исследована устойчивость их характеристик при отклонении законов распределения вероятностей выборок от нормального закона.

Введение. При обработке биомедицинских данных часто возникает задача сравнения показателей групп объектов (пациентов, животных или растений), принимавших тот или иной препарат или подвергавшихся какому-либо воздействию, и контрольных групп.

Наиболее часто требуется решить задачу о различии средних значений или о различии дисперсий (разброса значений показателей) по двум выборкам $x_1,..,x_n$ и $y_1,...,y_m$. Как правило, для формирования решения о средних значениях используют двухвыборочный *t*-критерий Стьюдента,

$$t_p = \frac{|\mu_1 - \mu_2|}{\sqrt{\sigma_1^2 + \sigma_2^2}}, \qquad (1)$$

где μ_1, μ_2 – соответственно выборочные математические ожидания первой и второй выборок; σ_1, σ_2 – их выборочные среднеквадратические отклонения; а для формирования решения о дисперсиях – *F*-критерий Фишера.

Критическая область, в которой гипотеза о равенстве дисперсий двух выборок отвергается при уровне значимости $\leq \alpha$, определяется как [2]

$$\left|\ln\frac{S_i}{S_k}\right| > \chi_{1-\frac{\alpha}{2}}, \qquad (2)$$

где S_i^2, S_k^2 – выборочные дисперсии *i*-й и *к*-й выборок соответственно; $\chi_{1-\frac{\alpha}{2}}$ – квантиль уровня $1-\frac{\alpha}{2}$ распределения Фишера.

Эти критерии являются оптимальными при условии нормального распределения показателей. При этом двухвыборочный *t*-критерий Стьюдента требует равенства дисперсий двух генеральных совокупностей, из которых формируются выборки.

Если реальные условия не соответствуют указанным условиям оптимальности, то распределения проверочных статистик (*t*-критерия Стьюдента и *F*-критерия Фишера) не подчиняются распределению Стьюдента и распределению Фишера соответственно. Эти обстоятельства приводят к неверным оценкам достоверности статистических выводов (вероятностей ошибок первого и второго рода). В результате может быть принято неверное решение об эффективности того или иного препарата или физиологической процедуры.

Постановка задачи. Опыт обработки реальных данных показывает, что в 50...80 % случаев количественные показатели биологических объектов не подчиняются нормальному распределению. Наиболее характерно это для опытных групп. По сравнению с группами контроля в опытных группах меняется форма распределения: затянуты левые или правые хвосты распределений, наблюдаются также значительные изменения дисперсии, асимметрии и эксцесса.

Проиллюстрируем сказанное на результатах следующих исследований. Проводились обследования группы лиц мужского пола в возрасте от 18 до 20 лет с целью определения

функциональных изменений состояния организма, которые могут произойти во времени. Обследования проводились с интервалом полгода. С точки зрения функциональной диагностики в качестве интегральной оценки динамики состояния организма определена динамика сердечного ритма, которая анализировалась с полугодовым интервалом регистрации

Обследования проводились при проведении активной ортопробы, которая состоит в том, что обследуемый находится в состоянии покоя сидя 10 минут, а затем встает и находится в состоянии орто (стоя) также 10 минут. Во время проведения пробы непрерывно записывается сигнал ЭКГ во втором стандартном отведении, и по этому сигналу программно строится функция сердечного ритма. Анализ сердечного ритма в соответствующих положениях тела – покоя (сидя) и орто (стоя) проводится по двух–пятиминутным интервалам сердечного ритма, в течение которых процесс является стационарным.

Реализации кардиоритма в положении стоя до проведения обследования и после показаны на рис. 1, 2, примеры их распределений – на рис. 3, 4.



Рис. 1. Кардиоритм в положении стоя:

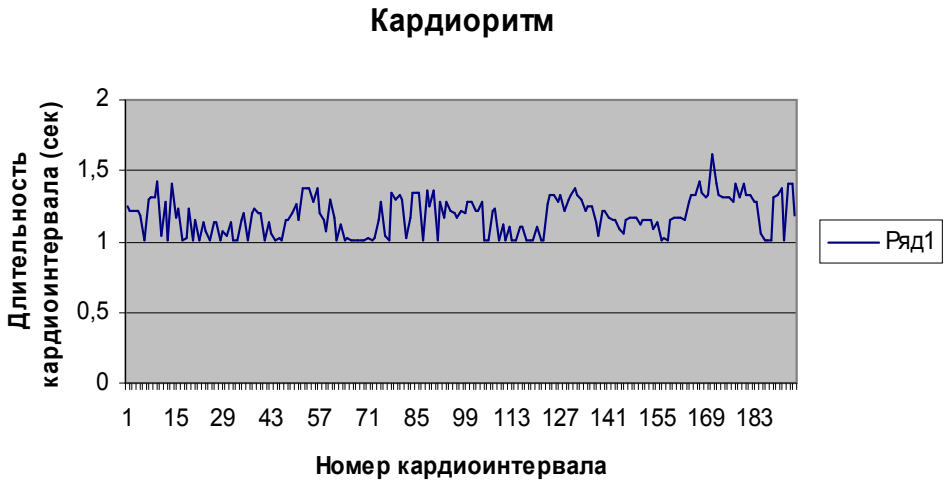


Рис. 2. Кардиоритм в положении стоя (через полгода)

Применение двухвыборочного *t*-критерия Стьюдента к решению задачи о различии двух распределений по этим данным приводит к решению об отсутствии различий, так как выборочные средние обеих выборок близки, хотя изменения в законе распределения вероятностей после воздействия очевидны (t_p = 0,097414, табличное значение критерия при уровне значимости 0,05 t = 1,96, t_p < t).

Аппроксимация гистограмм постоена в системе STATISTICA [3].

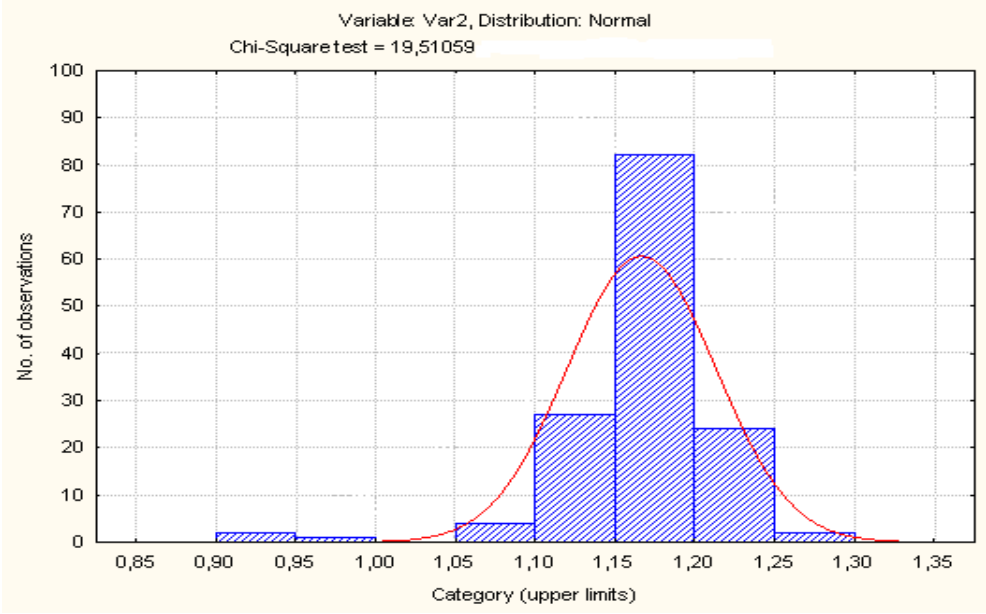


Рис. 3. Гистограмма кардиоритма в положении стоя (Chi-Square test – значение критерия χ^2)

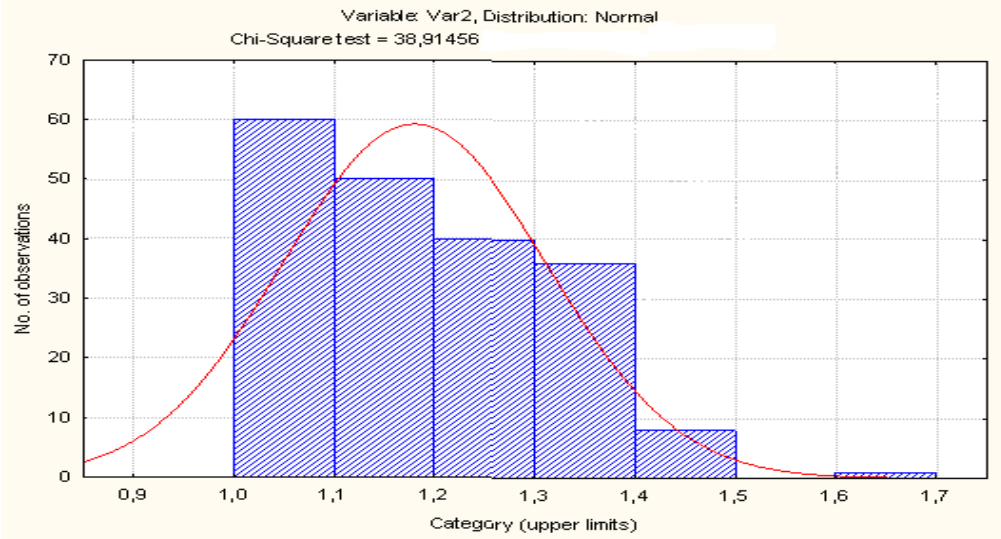


Рис. 4. Гистограмма кардиоритма в положении стоя (через полгода)

Проверка этих гистограмм на соответствие нормальному закону распределения вероятностей по критерию χ^2 не позволяет принять гипотезу о нормальном распределении признаков (поскольку вероятность $P(\chi^2 \geq \chi^2_q) = 0{,}004$ для первой реализации меньше уровня значимости 0,05, величина χ^2_q принимает большое значение для второй реализации) и, следовательно, применение *t*-критерия Стьюдента приведет к ошибкам.

В работе рассматривается задача применения непараметрического критерия для проверки гипотезы о равенстве распределений двух выборок, а также анализ его эффективности при отклонениях выборочных распределений от гауссовского.

Ранговый двухвыборочный критерий. Предлагаемый критерий основан на ранговой статистике и определяется следующей формулой:

$$S = \sum_{i=1}^{n} \left| R_i - \frac{m}{2} \right| > V_\alpha . \qquad (3)$$

Здесь R_i – ранг x_i в вариационном ряду, построенном по выборочным значениям $x_i, y_1,...,y_m$; V_α – порог принятия решения; α – вероятность ошибки первого рода (уровень значимости).

При гипотезе H_0: выборки принадлежат одной генеральной совокупности. Можно показать, что математическое ожидание и дисперсия статистики (3) определяются формулами:

$$m_1\{S\} = \frac{mn}{4}; \tag{4}$$

$$\mu_2\{S\} = \frac{mn(m+n+1)}{48}. \tag{5}$$

Порог принятия решения V_α с достаточной точностью можно определить по формуле

$$V_\alpha = m_1\{S\} + \chi_{1-\frac{\alpha}{2}}\sqrt{\mu_2\{S\}},$$

где $\chi_{1-\frac{\alpha}{2}}$ – квантиль уровня $1-\frac{\alpha}{2}$ стандартного нормального распределения.

По ранговому критерию (3) данные выборок различимы ($S > V_{0,05}$).

Сравнительный анализ эффективности алгоритмов. Рассмотрим задачу сравнительного исследования эффективности и устойчивости тестов (1), (2) и (3).

Введем параметр U, учитывающий силу воздействия (дозу препарата, физическую либо психологическую нагрузку).

Пусть доза U приводит к увеличению среднего значения измеряемой характеристики на pU и к увеличению дисперсии этой характеристики на

$$\frac{(1-p)U^2}{2},$$

где p – весовой коэффициент.

Построим кривые чувствительности трех алгоритмов (1), (2) и (3) к изменениям закона распределения измеряемой характеристики к воздействию U.

Пусть при отсутствии воздействия (контрольная группа) характеристика X распределена по нормальному закону с математическим ожиданием $m_1\{X\} = 0$ и дисперсией $\mu_2\{X\} = 1$, а воздействие U изменяет параметры распределения характеристики X таким образом, что

$$m_1\{X\} = pU;$$

$$\mu_2\{X\} = 1 + (1-p)\frac{U^2}{2}.$$

Для построения кривой чувствительности сформируем две выборки:

$$x_1,..,x_n \text{ и } y_1,...,y_m.$$

Выборку $x_1,..,x_n$ сформируем из нормированного нормального распределения по алгоритму

$$x_i = \sqrt{-2\ln(\eta_i)}\cos(2\pi\eta_{i+1}), i = \overline{1,n},$$

где η_i – значения случайной величины, распределенной по равномерному закону на интервале $[0,1]$.

Выборку $y_1,...,y_m$ сформируем из нормального распределения по алгоритму

$$y_i = \sqrt{-2(1+(1-p)\frac{U^2}{2}\ln(\eta_i)\cos(2\pi\eta_{i+1}) + pU}, i = \overline{1,m}.$$

Полученные выборки будем обрабатывать по алгоритмам (1), (2) и (3).

Для вынесения решения об обнаружении различия в распределениях (обнаружение эффекта от воздействия U) определим соответствующие пороги принятия решения как квантили уровня 0,05 распределения Стьюдента, распределения Фишера и для распределения ранговой статистики (3) по формулам (4) и (5). Кривые чувствительности методов или характеристики обнаружения строятся как зависимости вероятности превышения соответствующих порогов от величины воздействия U .

Семейство характеристик чувствительности исследуемых алгоритмов при разных значениях параметра p приведено на рис. 5. При этом использовались объемы выборок: $n = 5$, $m = 15$.

Из приведенных характеристик видно, что при заданной модели влияния воздействия (нагрузки U) на параметры распределения измеряемых характеристик при $p = 0,5$ чувствительность всех трех алгоритмов эквивалентна.

При $p > 0,5$ ранговый алгоритм уступает алгоритму Стьюдента (около 3 дБ при $p = 0,99$), однако превосходит чувствительность алгоритма Фишера

При $p < 0,5$ ранговый алгоритм превосходит по чувствительности алгоритм Стьюдента, но уступает алгоритму Фишера (около 3 дБ при $p = 0,01$)

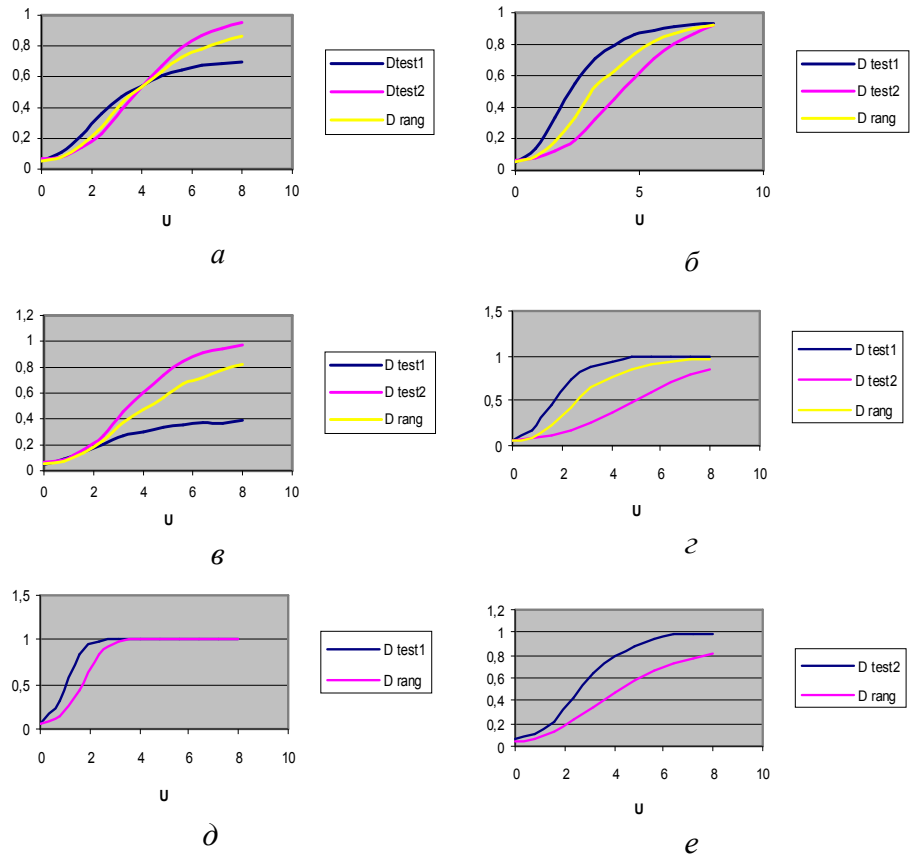


Рис. 5. Кривые чувствительности критериев Стьюдента, Фишера и рангового критерия (D test1, D test 2, D rang – вероятность превышения порога по критериям Стьюдента, Фишера и ранговому критерию соответственно, U – нагрузка): *а* – $p = 0,5$; *б* – $p = 0,6$; *в* – $p = 0,40$; *г* – $p = 0,7$; *д* – $p = 0,99$; *е* – $p = 0,01$

Исследование устойчивости характеристик рангового алгоритма. Для исследования устойчивости уровня значимости критериев (1 – 3) построим смесь гауссовских распределений.

$$f_{\varepsilon,p}(x) = \frac{(1-w)}{\sqrt{2\pi}}\exp\{-\frac{(x+\varepsilon)^2}{2}\} + \frac{w}{\sqrt{2\pi}}\exp\{-\frac{(x-\varepsilon)^2}{2}\}, \quad], w \in [0,1]. \qquad (6)$$

Это распределение зависит от двух параметров w и ε. Изменяя их можно придавать функции (6) различный вид меняя ее асимметрию и эксцесс. Распределения (6) существенно отличаются от нормального Различные варианты этого распределения приведены на рис. 6. Построение смеси (6) иллюстрируется рис. 6.

Кривые на рис. 7 и 8 показывают существенную зависимость уровня значимости критериев Стьюдента и Фишера от отклонения формы распределений признаков от нормального распределения и независимость этой характеристики рангового критерия.

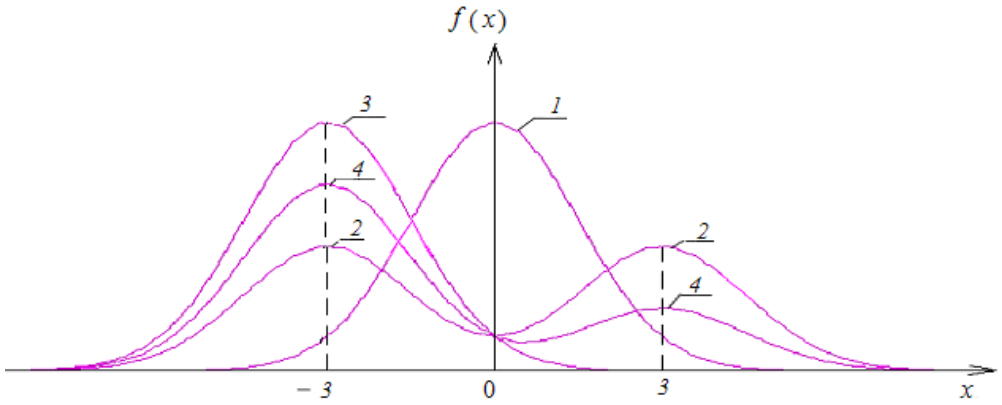


Рис. 6. Смеси гауссовских распределений (6) при различных параметрах w и ε
1 – $w = 0, \varepsilon = 0$; 2 – $w = 0{,}5, \varepsilon = 3$; 3 – $w = 0, \varepsilon = 3$; 4 – $w = 0{,}25, \varepsilon = 3$

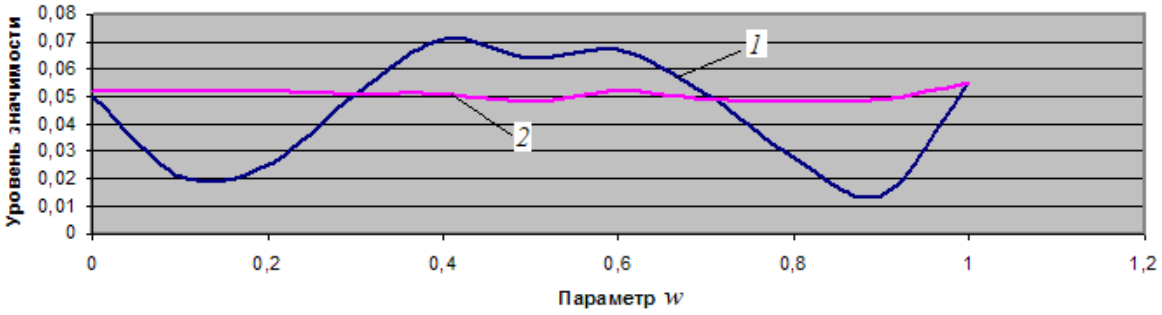


Рис. 7. Зависимость уровня значимости критерия Стьюдента (кривая *1*) и рангового критерия (кривая *2*) от параметра w при $\varepsilon = 3$

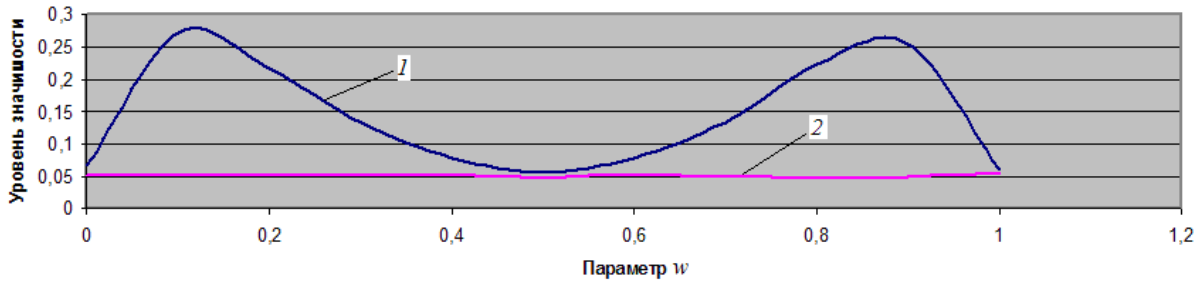


Рис. 8. Зависимость уровня значимости критерия Фишера (кривая *1*) и рангового критерия (кривая *2*) от параметра w при $\varepsilon = 3$

Заключение. Проведенные исследования позволяют сделать вывод о предпочтительности применения рангового критерия для задач проверки гипотезы о различии распределений двух выборок в условиях априорной неопределенности формы распределений.

Список литературы

1. *Я. Гаек*, *З. Шидак*. Теория ранговых критериев: Пер. с англ. – М.: Наука, 1971. – 376 с.
2. *Г. Корн*, *Т. Корн*. Справочник по математике для научных работников и инженеров. – М.: Наука, 1970. – 720 с.
3. *А. А. Халафян*. STATISTICA 6. Статистический анализ даннях: Учеб. 3-е изд. – М.: ООО "Бином-Пресс", 2007. – 512 с.

И. Г. Прокопенко, А. И. Чурина

Дослідження ефективності статистичних критеріїв в завданнях обробки біомедичних даних

Розглянуто завдання порівняльного аналізу ефективності трьох алгоритмів прийняття: критерія Стьюдента, критерія Фишера і двовибіркового рангового критерію. Досліджено стійкість їх характеристик у разі відхилення законів розподілу ймовірностей виборок від нормального закону.

I. G. Prokopenko, A. I. Churina

Research of statistical criteria efficiency in tasks of biomedical data processing

The comparative effectivity of three decision rules in the task concerning equality of distribution of two samples Student criterion, Fisher criterion and rank criterion, is considered. Robustness of their characteristics to deviations of the probability distribution laws from normal law is investigated.